DOCUMENT RESUME

ED 458 181 SO 033 111

AUTHOR Beins, Bernard C.
TITLE Writing in Statistics: Meaning Is Where You Find It.
PUB DATE 2001-08-24
NOTE 9p.; Paper presented at the Annual Meeting of the American Psychological Association (109th, San Francisco, CA, August 24-28, 2001).
PUB TYPE Opinion Papers (120) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Classroom Techniques; *Content Area Writing; Higher Education; Psychology; Social Science Research; *Statistics; Technical Writing; *Writing Assignments; *Writing Strategies
IDENTIFIERS *Data Sets; Writing Contexts

ABSTRACT

When students study the discipline of statistics, a domain that can be remote and abstract for them, it is critical that they understand what the numbers mean and how those numbers help people arrive at decisions. This paper presents different approaches that help students learn how researchers actually work with statistics and shows how students can develop skills by communicating statistical results to others. The paper proposes that writing about data enhances student learning. It outlines writing strategies to bring the domains of research and experience together. In the author/educator's classes, students use small data sets to provide descriptions of what the numbers tell them, interpreting the data with an eye to the importance of personal perspective and historical context in their conclusions. They assess larger and more complicated data sets and write essays about the results. The paper finds that these assignments allow students to (1) learn to translate numbers into comprehensible, nontechnical language; (2) show them how the same data set can provide data that could lead to conflicting interpretations; (3) help them learn how to identify appropriate uses of the data; (4) convince them that, although the numerical results of research may be incontrovertible, the meaning of those numbers emerges only when someone uses them to draw a conclusion; and (5) illustrate that the "truth" of the meaning of data is often dependent on perspective and historical context. Contains 3 notes, a figure, a table, 6 references, and a sample data set. (BT)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 458 181

Writing in Statistics: Meaning Is Where You Find It.

Beins, Bernard C.

SO 033 111

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Bernard C. Beins

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

Writing in Statistics: Meaning is Where You Find It[1]

Bernard C. Beins[2]
Ithaca College and American Psychological Association

Students learn best when they are actively engaged in the ideas they encounter. When they learn statistics, a domain that can be remote and abstract for them, it is critical that they understand what the numbers mean and how those numbers help us arrive at decisions. I present different approaches to helping students learn how researchers actually work with statistics and how students can develop skills communicating statistical results to others.

Students who learn statistics often develop a compartmentalized view of the area. They seem to view the world either as the domain of rejecting or failing to reject the null hypothesis (with its attendant and mysterious jargon) or as the domain of reality (which they believe they can understand). There is seldom a marriage of the two.

Fortunately, there are ways to enhance student learning, such as writing about data (Beins, 1993; Dunn, 2000) and applying it to life (Morgan, 2001; Weaver, 1992). There are also other illustrations of the connection between research and life (e. g., Beins, 1985).

The reason that these strategies increase student learning is that the students become engaged in issues of interest to them. In this presentation, I will outline writing strategies that will bring the domains of research and experience together.

In my classes students use small data sets to provide descriptions of what the numbers tell them. They interpret the data with an eye to the importance of personal perspective and historical context in the nature of their conclusions; they also assess larger and more complicated data sets and write essays about the results.

The effect of these assignments is to allow students (a) to learn to translate numbers into comprehensible and nontechnical language, (b) to show them how the same data set can provide data that could lead to conflicting interpretations, (c) to help them learn how to identify appropriate uses of the data, (d) to convince them that, although the numerical results of research may be incontrovertible, the meaning of those numbers emerges only when somebody uses them to draw a conclusion, and (e) to illustrate that the "truth" of the meaning of data is often dependent on perspective and historical context.

Students learn to connect data to decision making. They can move their conclusions out of the technical realm into areas of life that have meaning for them. As such, statistics emerges as a discipline of critical thinking about psychological topics.

Context and Data

It isn't immediately apparent to students (nor to many faculty) that research questions, and therefore the "truth", are dependent on the social context in which the research emerges.

---

[1]Presented at the annual convention of the American Psychological Association, San Francisco CA, Friday August 24, 2001.

[2]For further information, please contact me at the American Psychological Association, 750 First Street NE, Washington DC 20002-4242 (bbeins@apa.org) or at Ithaca College, Psychology Department, Ithaca NY 14850 (beins@ithaca.edu).

Another way to conceptualize this belief is that the data we collect don't have any meaning until we interpret them. Such an idea does not mean that the validity of our findings is evanescent; the idea merely implies that meaning requires an understanding of the context in which the data occur.

One of the issues facing the American public around the turn of the 20th century was whether women should sit on juries. There were numerous arguments against the practice, one of which came from Hugo Münsterberg, perhaps the most famous popular psychologist in the first decade of the 1900s.

Münsterberg performed an empirical study of women's suitability for jury duty. He wanted to know if women were able to debate in such a way as to improve the soundness of their judgments.

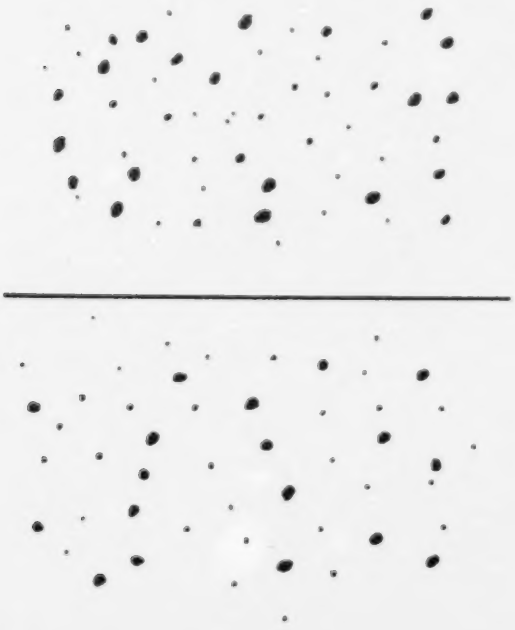


Figure 1. Stimuli based on Münsterberg's (1914) description of his jury research.

Using stimuli similar to those in Figure 1, he asked participants to judge which of the two displays contained more dots. After an initial judgment, the mock jurors discussed their views, then voted a second time. Although women and men were equally accurate at the beginning, essentially at chance levels, the accuracy of the men increased dramatically, to almost 80%, after discussion. The women's accuracy, about 50%, showed no such improvement.

Based on such a study, can you make a recommendation about whether our society gains from having female jurors? Table 1 presents some questions that will put the data and the research in perspective.

This simple study is easy to replicate. In reality, students don't need to discuss the issue of which stimulus has more dots. They can understand the data after a single exposure and judgment. (Incidentally, there are 54 dots in the upper segment of Figure 1, 50 dots in the lower.)

Table 1. Questions for Discussion of Münsterberg's Jury Research

1. Suspending your own beliefs, what do these data suggest regarding whether women can profitably engage in a debate involving an ambiguous "truth"?
2. How would you interpret these data, if you assume they are valid? That is, why might women show less accuracy than men after the debate?
3. Can you come up with an alternate interpretation that runs counter to your first interpretation?
4. Which interpretation is better?
5. Why isn't anybody doing research like this today?

It is important for students to know that the purpose of research is to answer questions

as well as possible. Any single experiment will answer some of the questions, but will leave others unresolved. In addition, any interpretation is susceptible to further testing. The final point here is that in some contexts, a particular research question will be meaningful; this one happens not to be for us in the year 2001.

Exploring Alternate Explanations

We can easily find data on teen births (e.g., http://www.childrends.org). Looking at the data, we see that there is a positive correlation between the average temperature in a state (measured at the state capital) and birth rates, $r = .74$.

Why should such a relationship exist? Students can postulate many causes, but one distinct possibility seems quite likely: the southeastern states tend to have high birth rates; they also tend to have sex education programs that focus on abstinence. If such a message is ineffective, birth rates will rise.

This data set is a good example of one that lends itself to initially plausible rival hypotheses. It lends itself to writing assignments that propose followup research.

Simplifying Complex Statistics

The jury study provided some very simple data that generate multiple interpretations. The next example involves data that are more complex and that require descriptive statistics if we are to understand them. The data in this example appear in the Appendix.[3] They are a compilation of federal figures that colleges and universities are required to file with the federal government.

As you can see for the first school listed, it is not always easy to characterize an institution as safe or unsafe. For this campus, some crime figures are relatively high, whereas some are relatively low. Before we can begin any assessment, though, we have to assess the relative position of each value in the distribution. This is where the mean and standard deviation are important. This type of data set also lends itself well to visual displays like boxplots.

The first school has a high rate of incidents involving alcohol and other drugs, but a low rate of violent crime. So does this mean that the school is safe or unsafe? It all depends on how you view each type of crime.

In my classes, I ask students to pick out their institution (Ithaca College) and write a lurid story for a supposed tabloid in which they convince me that the school is unsafe. Then I ask them to write an essay by the school's public relations office, depicting the safety of the campus.

After they have completed the writing assignments, I ask them how they were able to create two contradictory messages. It becomes pretty clear that you can slant your message by selecting some data and ignoring others. You haven't lied–exactly, but you haven't really told the truth, either.

In a final essay, the students present a balanced picture of the crime situation in a readable and interesting way. They aren't supposed to list a set of numbers (pretty boring), but they are to paint an accurate picture.

Even in the final essay, students have to select some information to highlight and some to ignore. This is an important point about analyzing data: you have to figure out which numbers are more important and which you can safely ignore.

[3]These data were compiled several years ago; I removed the names of the institutions for reasons of courtesy to the institutions. Data are available at <http://ope.ed.gov/security/index.asp>.

Finding Useful Data

There are typically too many numbers in our lives to keep track of. The good news is that many of them are amenable to teaching about statistical ways of thinking.

The federal government provides a treasure trove of statistics on a wide variety of topics. The *National Center for Educational Statistics* <http://nces.ed.gov/> has compiled myriad statistics on topics that can interest students. *Child Trends* <http://www.childtrends.org/HomePg.asp>, a research organization that studies children, youth, and families through research, data collection, and data analysis, provides useful data as well.

The *General Social Survey* <http://www.icpsr.umich.edu/GSS/> has also published interesting data. You can use the easily accessible data to generate topics for discussion. After that, students can write essays in which they use data to make decisions and to support an argument.

For topics of immediate interest, you can go to the Gallup <http://www.gallup.com/> organization, which gives you access to a wealth of survey data.

Conclusion

The wealth of data at our fingertips (or our keyboards) has provided the opportunity to present our students with information about topics that interest them (and us). With access to these data, we can plan writing exercises that will let our students explore the meaning of data and then communicate what they have learned to us.

Depending on the course, the students, and the desired outcome, we can use these sources to help our students learn about descriptive and inferential statistics. We can also enhance their appreciation of how we use numbers to understand the behaviors of people around us.

References

Beins, B.C. (1985). Teaching the relevance of statistics through consumer-oriented research. *Teaching of Psychology, 12*, 168-169.

Beins, B. C. (1993). Writing assignments in Statistics Classes Encourage Students to Learn Interpretation. *Teaching of Psychology, 20*, 161-164.

Dunn, D. S. (2000). Letter exchanges on statistics and research methods: Writing, responding, and learning. *Teaching of Psychology, 27*, 128-130.

Morgan, B. L. (2001). Statistically lively uses for obituaries. *Teaching of Psychology, 28*, 56-58.

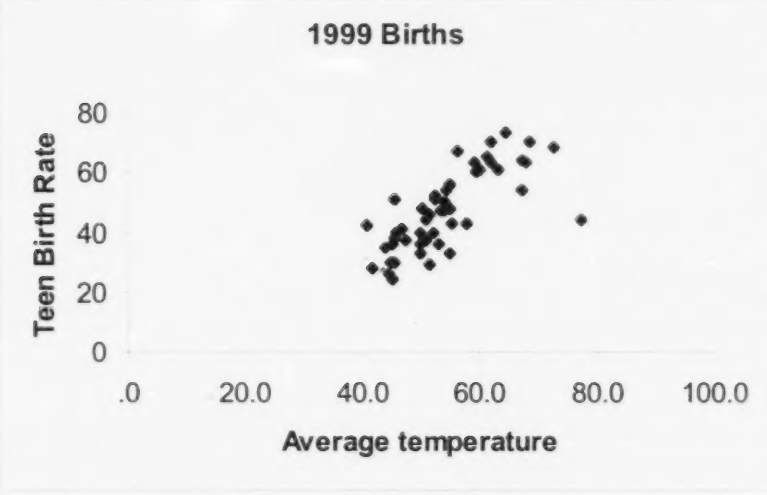
Münsterberg, H. (1914). *Psychology and social sanity*. NY: Doubleday, Page & Co.

Weaver, K. A. (1992). Elaborating selected statistical concepts with common experience. *Teaching of Psychology, 19*, 178-179.

Appendix: Data set on campus crime.

Locale	Per capita crime	Total index	Violent Crimes	Assaults	Theft	Drugs	Vandalism	Enrollment
Urban	1	8	0	0	7	2	12	9716
Urban	9	177	7	6	170	0	9	20760
Rural	31	191	2	0	183	92	238	6105
Urban	20	73	0	0	73	57	48	3602
Urban	1	3	2	1	1		2	4492
Urban	6	20	1	2	19	14	5	3231
Urban	29	549	5	2	544	0	47	18992
Rural	56	639	11	10	626	9	181	11319
Urban	12	156	1	11	152	9	44	13036
Urban	4	26	2	2	23	0	12	6049
Urban	17	68	1	8	64	65	66	4016
Urban	19	58	7	15	575	99	38	30753
Urban	11	181	2	23	177	23	194	16465
Urban	10	35	3	1	27	0	0	3558
Rural	56	355	13	28	342	4	115	6371
Urban	37	463	0	112	462	8	541	12494
Rural	33	115	0	0	115	91	143	3481
Urban	31	511	7	66	501	16	340	16500
Urban	22	282	2	27	280	15	241	12588
Urban	27	242	2	37	239	9	150	8840
Urban	27	746	9	69	717	10	377	28005
Rural	27	193	2	25	191	4	192	7263
Rural	33	163	0	14	163	14	93	4994
Rural	44	234	1	6	231	5	331	5321
Urban	17	140	2	22	138	9	133	8097
Urban	15	61	4	12	57	1	57	3948
Rural	22	134	3	14	131	34	148	6017
Rural	29	254	1	27	252	30	194	8672
Rural	17	99	1	7	98	1	97	5958
Rural	29	108	1	6	107	1	74	3696
Urban	39	163	1	0	161	5	125	4144
Urban	51	861	13	75	837	2	485	16747
Urban	16	198	7	24	190	10		12721
Urban	19	204	9	42	191	4	135	10802
Urban	29	634	19	14	612		452	22085
Urban	1	5	0	0	4	0	2	3722
Urban	74	641	5	1	612			8634
Mean	24.9	243.0	3.9	19.2	250.6	18.9	152.0	10086.3
St. Dev.	16.5	228.7	4.5	25.1	227.0	28.1	145.4	7041.6
Locale	**Per capita crime**	**Total index**	**Violent Crimes**	**Assaults**	**Theft**	**Drugs**	**Vandalism**	**Enrollment**

There is a predictable pattern of teen births in this country. The higher the temperature, the higher the likelihood that teenagers will become pregnant ($r = .741, p < .01$).



Are there more brown dots (on top) or more blue dots (on the bottom)?